

Analysis of the BEAVRS benchmark using the GPU-based direct whole core calculation code CRANE

Authors: Li, Mr. Jun, Chen, Dr. GuoHua, Zhang, Prof. Qian, RuoHan Zheng, LiQin Zhang, Zhang, Prof. Shaohong, Zhang, Prof. Shaohong

Date: 2025-05-09T22:58:54+00:00

Abstract

The BEAVRS (Benchmark for Evaluation and Validation of Reactor Simulation) benchmark proposed by the Massachusetts Institute of Technology (MIT) computational reactor physics group is widely used by various institutions in the world to verify and validate their new generation direct whole core calculation codes. It enables analysts to develop an extremely detailed PWR core model and to carry out multi-physics coupled verification and validation for the reactor core analysis code. Although there are numerous publications in the literature reporting the validation results, few publish thorough validation results against all the measurement data of the problem. In this paper, the BEAVRS benchmark is solved using the CRANE code, which is a GPU-based deterministic direct whole core calculation code for PWR. It is demonstrated that direct whole core calculation with detailed core model can be performed on a mini server mounted with 10 consumer-grade RTX 3090 graphics cards, and the average time needed to complete neutronics and thermal-hydraulics coupled analysis for a single core state can be shortened to a record-breaking about one minute. Thorough verification and validation against all the measurement data shows that the predicted criticality, control rod bank worths, in-core detector signal distribution and the boron let-down curves of two cycles agree well with the measurements. These results indicate that even by exploiting the computing power of today's consumer-grade GPUs, direct whole core calculations with detailed core model for large commercial PWRs are now feasible, and CRANE is ready for PWR practical applications in terms of both solution fidelity and speed performance.

Full Text

Preamble

Analysis of the BEAVRS benchmark using the GPU-based direct whole core calculation code CRANE

Jun Li,¹ Guo-Hua Chen,^{2,3} Qian Zhang,¹ Ruo-Han Zheng,⁴ Li-Qin Zhang,⁴ and Shao-Hong Zhang^{1,†}

¹School of Physics, Zhejiang University, Hangzhou, 310027, China ²Shanghai NuStar Nuclear Power Technology Co. Ltd, Shanghai, 201101, China ³Advanced Technology Institute, Zhejiang University, Hangzhou, 310027, China ⁴Nuclear Power Institute of China, Chengdu, 610041, China

The BEAVRS (Benchmark for Evaluation and Validation of Reactor Simulation) benchmark proposed by the Massachusetts Institute of Technology (MIT) computational reactor physics group is widely used by various institutions worldwide to verify and validate their new generation direct whole core calculation codes. The benchmark enables analysts to develop an extremely detailed PWR core model and to carry out multi-physics coupled verification and validation for reactor core analysis codes. Although numerous publications in the literature report validation results, few publish thorough validation results against all the measurement data of the problem.

In this paper, the BEAVRS benchmark is solved using the CRANE code, which is a GPU-based deterministic direct whole core calculation code for PWR. It is demonstrated that direct whole core calculation with detailed core model can be performed on a mini server mounted with 10 consumer-grade RTX 3090 graphics cards, and the average time needed to complete neutronics and thermal-hydraulics coupled analysis for a single core state can be shortened to a record-breaking about one minute. Thorough verification and validation against all the measurement data shows that the predicted criticality, control rod bank worths, in-core detector signal distribution, and the boron let-down curves of two cycles agree well with the measurements.

These results indicate that even by exploiting the computing power of today's consumer-grade GPUs, direct whole core calculations with detailed core model for large commercial PWRs are now feasible, and CRANE is ready for PWR practical applications in terms of both solution fidelity and speed performance.

Keywords: BEAVRS benchmark, direct whole core calculation, CRANE, GPU computation

Introduction

A nuclear reactor is a typical complex system with multi-physics coupling. Due to the high cost of conducting nuclear reactor experimental research, numerical simulation has always been an important means of investigating the behavior

of nuclear reactors. With the rapid development of computer hardware performance, since the beginning of this century, researchers have made significant progress in conducting first-principle-based simulations of the multi-physics behaviors in nuclear reactors. A batch of new generation codes have been successfully developed, adopting either the deterministic approach or the probabilistic approach. For instance, codes DeCART[1], nTRACER[2], MPACT[3], and NECP-X[4] are the deterministic ones primarily based on the method of characteristics (MOC), while codes MC21[5], RMC[6], OpenMC[7], JMCT[8], and SuperMC[9] are the probabilistic ones employing the Monte Carlo (MC) method. All these codes are able to perform high-fidelity direct whole core calculations for PWRs without introducing any empirical hypothesis as in the conventional two-step analysis method.

BEAVRS[10] is a benchmark problem proposed by the MIT Computational Reactor Physics Group in 2013 to meet the needs of verifying and validating the new generation of reactor core analysis codes. This problem is based on a 4-loop pressurized water reactor (PWR) designed by Westinghouse in the United States. It provides highly detailed as-is design data for the reactor and measurement data for the initial two operating cycles to the public to allow for validation of method developments. It enables analysts to develop an extremely detailed PWR core model and carry out multi-physics coupled verification and validation for the reactor core analysis code. Since previous light water reactor benchmarks can only be used to verify a single physical problem, and most institutions cannot have access to this kind of utility proprietary information, this problem is unique and valuable.

Considering the complexity of the problem and the computation cost that it would take to perform direct whole core multi-physics coupled calculation with very detailed core modeling, it is a challenging task to complete the high-fidelity analysis of the BEAVRS benchmark problem. So far, various institutions worldwide have attempted to solve this problem using their new generation codes. However, the degree of completion varies among the institutions. Some have only completed the calculation for the hot zero power (HZP) state at the beginning of the first cycle (BOC1)[11–16], while others have completed both the BOC1 HZP state calculation and the first cycle core follow calculations[17–22]. Only a few institutions have completed all the two cycles calculation provided in the benchmark[23–27]. As for the validation of computational results against the measurement data, most institutions have conducted comparisons of BOC1 HZP critical boron concentration (CBC) and the boron letdown curve for the initial cycle, while only a few have carried out comparisons of control rod bank worth and in-core flux measurement results at different burnup of the reactor. So far, there is no literature published yet reporting the comparison of control rod bank worth between the prediction and the measurement data obtained at the low-power physics test (LPPT) stage at the beginning of cycle 2 (BOC2).

This work focuses on using the GPU-based deterministic direct whole core calculation code CRANE[28] to analyze the BEAVRS benchmark for assessing its

solution fidelity and speed performance for PWR practical applications. It is demonstrated that just by exploiting the computing power of today's consumer-grade GPUs, direct whole core calculation with detailed core model and reactor core flow calculation adopting the exact reactor power operation history are feasible. Thorough verification and validation of CRANE-predicted results against the measurement data of the benchmark are performed, and parameters such as the control rod worth (CRW) at BOC2, which have never been reported in previous literature, are presented with comparative results in this work.

The rest of the paper is organized as follows. Section II outlines the methodology that CRANE adopted for PWR neutronics, thermal-hydraulics (TH), and fuel rod performance coupled calculation. Section III gives the necessary details about the CRANE BEAVRS core modeling. Section IV presents the validation results against all the measurement data from the two operating cycles of the problem, and also the comparisons of prediction accuracy of CRANE with that of relevant codes in the world. Section V gives the introduction of the low-cost CPU-GPU heterogeneous platform adopted for this study and various measures the code developers used to enhance software performance, while Section VI concludes the paper.

II. CRANE Methodology

CRANE is an advanced deterministic direct whole core calculation code for PWRs. It has been developed in Shanghai NuStar Nuclear Power Technology Co., Ltd for more than 5 years as its next generation core analysis code targeting commercial PWR applications. The code is able to perform high-fidelity reactor neutronics/TH coupled analysis. The main methodology implemented in the code is outlined in this section.

The neutronics module of CRANE has extensively inherited many methods and good practices from NuStar's 2D lattice physics code ROBIN[29], such as the 69-group cross-section data library processed from the ENDF/B-VII.1 evaluated nuclear data file, the resonance calculation method applicable to irregular geometries that integrates the traditional equivalence theory and the enhanced neutron current method, and the use of pre-calculated correction factor tables to account for the resonance interference effects among major resonant nuclei. To achieve a three-dimensional neutron transport solution for the entire reactor with fine resolution down to the sub-pin level, the 2D/1D coupled iteration method is adopted. MOC is employed to solve both the 2D planar problem and the axial 1D problem. Iteration acceleration is achieved by successively solving the pin-level multi-group and few-group partial current-based 3D Coarse Mesh Finite Difference (pCMFD) problem. Implicit trapezoidal method and Chebyshev rational approximation (CRAM) are used to solve the depletion problem, and the strategies of linear reaction rate method (LR) and logarithmic linear reaction rate method (LLR) are used to deal with the coupling calculation of neutron transport and depletion for normal fuel pins and Gd-bearing fuel pins respectively.

The thermal-hydraulic module of CRANE includes a subchannel analysis model for calculating the temperature and density of the coolant, as well as a fuel rod analysis model for calculating the temperature of the fuel pellet. The subchannel analysis model is developed with reference to the CTF code[30]; it gives the temperature and density of the coolant around the fuel rod by solving the mass, energy, and momentum conservation equations for the two-fluid, three-field (i.e., fluid film, fluid drops, and vapor) system within the pin-level subchannel control volumes. The fuel rod analysis model is developed by employing the fuel performance analysis model that the FuSPAC code[31] used. Fuel rod behaviors under irradiation conditions caused by burnup degradation, swelling, expansion, creep, and stress-strain are accurately simulated, and the evolution of intra-pellet temperature distribution with burnup is generated.

It is noteworthy that as CRANE was developed entirely from scratch, the code developers optimized the algorithms and data structures of different physical modules from the very beginning, achieving high-efficiency multi-physics coupling iterations through integrated code development and compilation. For instance, unlike conventional approaches that rely on external file exchanges or dynamic link library interfaces for multi-physics coupling, CRANE implements direct memory access for data sharing between modules. This approach not only prevents memory waste caused by duplicate parameter allocation across different program modules but also eliminates the additional communication overhead inherent in traditional data-sharing methods, thereby facilitating efficient multi-physics coupling iterations.

Another example involves mesh partitioning: while different physics modules require distinct mesh configurations due to their specific solving requirements—neutronics calculations demand fuel-rod-centered meshes whereas subchannel thermal-hydraulics analysis requires coolant-centered meshes—CRANE establishes a unified foundational mesh during modeling that accommodates all physical modules' needs. Based on this primary mesh, the system creates fixed mapping relationships for meshes and data across different physics modules, thereby maximizing the efficiency of multi-physics coupling iterations.

Moreover, it is worthy emphasizing that unlike most deterministic direct whole core calculation codes that run on dedicated CPU-based supercomputers, CRANE by design is a fully-GPU-based code and its target computing platform is chosen as an industry-cost-effective computer server mounted with multiple consumer-grade GPUs. As depicted below in Fig. 1 [Figure 1: see original paper], the flow chart of CRANE shows modules marked in green running on GPU and modules marked in blue running on CPU. Fig. 2 [Figure 2: see original paper] and Fig. 3 [Figure 3: see original paper] show the core model that CRANE outputted by exploiting the Python visualization tool package.

III. Core Modeling

The reactor core specified in the BEAVRS benchmark problem is a four-loop Westinghouse PWR loaded with 193 optimized fuel assemblies of 17×17 lattice for the rated power of 3411 MWth. The benchmark specification provides all the detailed geometric dimensions and material compositions for the major core constituents, including fuel assemblies with different fuel enrichment, borosilicate glass burnable absorber, control rods, core baffle, and barrel. As the first step of the benchmark calculation, a virtual core model that faithfully represents the practical core is set up. Since CRANE employs Constructive Solid Geometry (CSG) method for modeling in the x-y plane, it is capable of faithfully modeling all the geometrical details of the problem without introducing any approximations.

More detailed information about CRANE BEAVRS core modeling is given as follows. In the radial direction, the 1/4 symmetric core is modeled, with the outer boundary set at one fuel assembly width outside the active core. In the axial direction, the code automatically generates axial meshes according to material heterogeneity, with the maximum height of an axial layer set to be 20 cm. In the end, the total height of the problem is divided into 38 layers, with 29 of them located in the active core. There are 9 spacer grids in total along the height; each height containing the spacer grid is set as a separate layer, and all the spacer grid straps in these layers are explicitly modeled.

In terms of spatial mesh division for the neutronics calculation, Fig. 4 Figure 4: see original paper and (b) show the flat source approximation mesh division schemes that CRANE adopted for a normal fuel pin cell and a burnable absorber pin cell, respectively. The ray tracing parameters used for MOC calculation are as follows: for the radial 2D MOC calculation, the Tabuchi-Yamamoto optimum polar angle quadrature set[32] is employed, using 3 polar angles and 12 azimuthal angles for the octant of the solid angle sphere, and 0.05 cm ray spacing. For the axial 1D MOC calculation, the Legendre quadrature set is employed, using 8 discrete angles in the range of $0-\pi$ and 0.5 cm mesh height for flat source approximation.

Pin-wise thermal-hydraulics feedback is considered by solving the corresponding model for coolant and fuel rod; however, only the simple closed channel TH module is used instead of the more time-consuming elaborated sub-channel module, since the cross flow between adjacent pin cells is small and has a negligible effect on core reactivity and power distributions. Convergence criteria are set as follows: the criterion for critical boron concentration search is 0.5 ppm, the convergence criterion for calculating CRW and temperature coefficient is 0.2 pcm for k-effective, the convergence criterion for fission source is 5×10^{-4} , and the convergence criteria for TH feedback calculations are 0.2 K for coolant temperature and 1 K for fuel pellet temperature, respectively.

IV. Validation Results

1. HZP Core Calculation of Cycle 1

Table 1 shows the CBCs calculated by CRANE under 5 different control rod insertion states at BOC1 HZP state and the corresponding deviation from the measurements. The maximum deviation is 25 ppm, which meets well the requirement of American Nuclear Society (ANS) standards[33], i.e., the deviations between measured value and predicted value should be within ± 50 ppm. Reference[34] gives the standard deviation of the measured CBC, which is 22 ppm for all the 5 measurements. It can be seen that except for the D-in and CD-in states, where the deviations slightly exceed the standard deviation of the measurements, all the other 3 deviations lie well within the range of the standard deviation of the measurement.

To compare the prediction accuracy of CRANE with that of relevant codes in the world, Table 1 also presents the prediction deviation of the deterministic code VERA[27] and Monte Carlo code RMC[25]. The accuracy of CRANE for criticality prediction is comparable to that of VERA and RMC.

Table 2 shows the validation results of CRANE CRW prediction, where the measurement and the corresponding measurement errors are obtained from reference[34]. The maximum relative deviation for the worth of individual bank is -10%, and the deviation for the total integral of predicted worths is -1.5%. The CRANE prediction accuracy meets the requirements of the ANS standards[33], i.e., the deviation for the worth of individual bank should be within $\pm 15\%$ or 100 pcm, whichever is greater, and the deviation for the total integral of control rod worths should be within $\pm 10\%$ (for Dynamic Rod Worth Measurement, the total worth should be within 8%).

Similar to the comparison of the CBC, Table 2 also presents the prediction deviation of code VERA[27] and RMC[25]. It can be observed that CRANE's results demonstrate relatively closer agreement with that of RMC. A detailed examination of CRANE's prediction results reveals noticeably larger deviations in the B, A, and SE rod banks, with bank B demonstrating positive deviation while others show negative deviations. It is believed that this phenomenon stems from two primary causes: firstly, the inherent small worth values of bank A and SE introduce greater relative measurement errors[34]; secondly, the CRANE-predicted radial thermal neutron flux distribution at BOC1 state exhibits an outward-high/inward-low tilt compared to the measurements (see Fig. 8 [Figure 8: see original paper]). This flux distribution discrepancy results in systematic overestimation of peripheral bank B rod worth and underestimation of interior bank A/SE rod worths.

Table 3 shows the validation results for isothermal temperature coefficient (ITC). The benchmark gives 3 measured values at different control rod insertion states. For each core state, the CRANE prediction meets well the ANS standard requirements, i.e., the deviation should be less than ± 2 pcm/ $^{\circ}$ F.

2. Core Follow Calculation of Cycle 1

Fig. 5 [Figure 5: see original paper] gives the detailed power operation history of the initial cycle of BEAVRS benchmark, where the blue dots denote the times when the in-core neutron flux measurements are carried out. There are in total 24 in-core measurements performed during the whole cycle. For each in-core measurement, the uniqueness of this work is that it adopts the exact power history to perform core follow calculation, since today's GPU computing power enables CRANE to do so.

Fig. 6 [Figure 6: see original paper] gives the comparison of CBC obtained from the core follow calculation with those specified by the problem for 24 different burnup states, where the in-core flux measurement is carried out. The CRANE prediction agrees well with the measurement, and the deviations throughout the entire cycle are typically maintained within ± 25 ppm. The benchmark provides the reference boron letdown data for Cycle 1. In order to perform the boron letdown curve comparison, one has to obtain the CBC under the standard boron condition, i.e., the concentration under the all-rod-out (ARO) and hot-full-power (HFP) state. Although with the previous core follow calculation results available, this value can be determined by applying theoretical corrections for power and control rod effects, in this study, a more straightforward method is applied: directly performing the ARO HFP depletion calculation for Cycle 1.

Fig. 7 [Figure 7: see original paper] shows the obtained results with the reference ones, where the non-physical minus boron concentration at the end of cycle in the predicted value arises from the predicted cycle length being slightly shorter than the reference one. One may notice that for the whole core lifetime of Cycle 1, the deviation of standard boron concentration lies within the range of ± 25 ppm, which is consistent with the above-mentioned core follow calculation results. As a supplement, Fig. 7 also shows the boron letdown curve predicted by the nTRACER[23] code; the results of CRANE and nTRACER are also in good agreement. Both CRANE and nTRACER slightly underestimate the cycle length of Cycle 1, and the situation of CRANE is slightly better than that of nTRACER. Reference [34] investigated the uncertainty of BEAVRS CBC measurement by analyzing the results of multiple measurements of boron concentration on the same day and comparing the deviation between the measured values and the calculation results from the analysis code currently employed in engineering practice, and the conclusion was that the uncertainty is ± 25 ppm, which is entirely consistent with the deviations that one may observe in Fig. 6 and Fig. 7. Therefore, it is concluded that CRANE is able to accurately predict the BEAVRS core reactivity variation for the entire Cycle 1.

In-Core Detector Signal Calculation of Cycle 1

In this subsection, the accuracy of CRANE in calculating the 3D neutron flux distribution within the reactor core is validated against the in-core flux measure-

ments provided by the BEAVRS benchmark. Since the benchmark provides only the normalized 3D detector response, not the detector current signal itself, the in-core detector signal is simulated in this study by putting a trace amount of U-235 in the instrumentation thimble to generate the fission rates at all the detector locations. The resulting detector fission rates are then normalized to generate a 3D distribution that can be compared with the measured 3D detector response.

As shown in Fig. 5, many in-core flux measurements of Cycle 1 are carried out during the power ascension process. Since a change in power level induces delayed variations in xenon concentration and neutron flux spatial distribution within the reactor, and such transient fluctuations cannot be accurately calculated according to the operation history in days given by the benchmark, only 14 sets of measurement data obtained at relatively high-power level and at relatively stable states are selected to be compared with, just like the practice adopted in reference[27]. The comparison results are summarized in Table 4, where 2D RMS means the root mean square deviation between the predicted and measured 2D detector signal distributions. The 2D predicted and measured distributions are obtained by integrating the corresponding 3D detector signal distributions in the axial direction. The comparison given in Table 4 indicates that the accuracy of in-core detector signal prediction at different burnup for both CRANE and VERA are basically comparable.

To learn more about the detailed deviations in the detector signal distribution, Fig. 8 through Fig. 10 [Figure 10: see original paper] show 2D detector signal distribution deviations at representative burnup points (beginning, middle, and end of cycle). From these comparisons, it can be seen that in general the accuracy of CRANE prediction is satisfactory, with the maximum relative deviation being less than 5%. However, it can also be observed that compared to the measured signals, the CRANE prediction results exhibit a slight outward/inward tilt, which diminishes with increasing fuel burnup feedback.

4. HZP Core Calculation of Cycle 2

At BOC1, all the fuel loaded into the core is fresh fuel. For Cycle 2, which is a reloaded core, both depleted fuel and fresh fuel are loaded into the core at the beginning of the cycle. Engineering functions such as depleted fuel shuffling and removing the borosilicate glass burnable absorber rod from the depleted fuel from Cycle 1 are necessary for a code to perform Cycle 2 calculations. Therefore, the Cycle 2 problem provides not only a case to further validate the core depletion functionality of a core analysis code but also a good benchmark to check whether the code possesses the necessary functionality for engineering applications.

The CRANE code is adopted to perform HZP core calculation of Cycle 2 of BEAVRS benchmark. The CBC, ITC, and CRW corresponding to the measurements are predicted, and the comparison results are summarized in Table 5 and

Table 6 respectively. It can be seen that the predicted CBC and ITC are in good agreement with the measurements, and the degree of deviation between the prediction and measurement is similar to that of the first cycle. Meanwhile, it can also be seen that CRANE and VERA have comparable prediction accuracy for criticality calculation.

When looking at the validation results of CRANE CRW prediction, it is noticed that all the prediction accuracy meets the requirements of the ANS standard. However, unlike the case of the first cycle, there are now more banks of rod with deviations reaching around 10%, and the maximum relative deviation has reached 12.7%. Considering that the CRANE-predicted radial thermal neutron flux distribution at BOC2 no longer exhibits significant outward/inward tilt (see Fig. 14 [Figure 14: see original paper]), the reason for this phenomenon, as believed in this study, is mainly due to the errors in the measurements themselves. Comparing the measurement errors given in Table 6 and Table 2, one may find that the error is further increased at Cycle 2, especially for banks with relatively small CRW, such as bank SD, SC, and SA, where the measurement error itself exceeds the deviation. As mentioned earlier, there is no existing literature published on the CRW prediction for Cycle 2 of BEAVRS benchmark; therefore, there are no validation results of other codes provided in Table 6.

5. Core Follow Calculation of Cycle 2

In accordance with Cycle 1 practice, the Cycle 2 core follow calculation utilizes the exact power history shown in Fig. 11 [Figure 11: see original paper], and the CRANE predicted CBCs are compared against the measurements specified for 14 burnup states where the in-core flux measurement is carried out. Moreover, ARO HFP core depletion calculations are also performed, and the obtained boron letdown curve is compared against the reference one and the one predicted by the nTRACER code. Comparison results are shown in Fig. 12 [Figure 12: see original paper] and Fig. 13 [Figure 13: see original paper] respectively. It can be seen that except for one isolated instance of core follow calculation where the deviation exceeds 35 ppm, all the other deviations between the CRANE-predicted and the measured CBC are less than 25 ppm throughout the entire cycle. Considering that the uncertainty of BEAVRS boron concentration measurement is ± 25 ppm, it can be concluded once again that CRANE can predict the BEAVRS core reactivity variation for the reload cycle with satisfactory accuracy. The comparable accuracy between CRANE and nTRACER is also demonstrated.

In-Core Detector Signal Calculation of Cycle 2 Cycle 2 includes 14 in-core flux measurements. CRANE-predicted 2D in-core detector signal distributions are compared against the measurements, and the results are summarized in Table 7. To facilitate the performance assessment between CRANE and other high-fidelity codes, the in-core detector signal validation results of the deterministic code VERA and the Monte Carlo/TH coupled system (MCS+CTF)[24] are

also provided in Table 7. Based on these comparison results, it can be concluded that CRANE is able to accurately predict the in-core flux distribution and its variation with fuel burnup for the whole cycle. The accuracy of CRANE is fully comparable to that of VERA and generally better than that of the MCS/CTF coupling system. Nevertheless, unlike the VERA case where the RMS errors at the first two burnup points are evidently larger, the errors of CRANE are rather stable throughout the cycle.

Similar to the case of Cycle 1, three representative detector signal deviation distributions are shown in Fig. 14 to Fig. 16 [Figure 16: see original paper]. Compared with the ones for Cycle 1, the results of Cycle 2 are overall better; the relative deviation for all the results is within 5%, and the noticeable outward/inward tilt that existed in the BOC result for Cycle 1 no longer exists for this cycle.

V. Computing Platform and Code Performance

As the core analysis method shifts from the conventional two-step procedure to the direct whole core calculation involving first-principle-based simulations of the multi-physics behaviors in nuclear reactors, the computation load increases significantly. It is estimated that the computation load of the new method is at least 2 to 3 orders of magnitude greater than that of the conventional method. Desktop personal computers or workstations, which are the computing platforms for the conventional method, are no longer suitable for the new method due to their limited computing power and memory. High-performance computing (HPC) clusters or supercomputers equipped with many CPUs are chosen by many institutions worldwide as the computing platform for the direct whole core calculation codes. However, due to the complexity of the problem, even by exploiting the computing power of today's HPC clusters or supercomputers, the computing time of the direct whole core calculation code still poses an issue for its use in routine design analysis. For instance, to simulate the BEAVRS benchmark, VERA takes 50 minutes (760 CPU-hours) on average to complete one state core analysis on a cluster with 880 2.3 GHz Intel Xeon processors with 4 GB RAM[27], while nTRACER takes 3 hours on a cluster with 28 computing nodes mounted with dual 2.67 GHz Intel Xeon X5650 processors yielding 12 cores per node[23]. Therefore, to reduce the computing time of a new generation code to a level acceptable for practical engineering applications, supercomputers mounted with thousands of cores are necessary. Unfortunately, such large-scale HPC clusters are usually available only at national laboratories, not at nuclear design institutes, not to mention ordinary research groups. Both the computing time and platform availability issues thus pose key obstacles that hinder the widespread use of these new generation codes.

Meanwhile, in the past two decades, the computing power of GPUs has exhibited much stronger growth than that of CPUs. The computing power that used to be provided by multiple CPU nodes can now be provided by a single GPU card. GPUs provide much higher computing power than CPUs for the

same energy consumption. It thus provides researchers a new powerful platform to consider for the development of direct whole core calculation codes for PWRs, i.e., adopting heterogeneous computing that utilizes more GPUs than CPUs. The past decade has witnessed growing adoption of GPU acceleration in neutron transport computation. Starting with initial test applications[35, 36], the technology has progressively expanded to wider applications in both deterministic neutron transport methods[28, 37–44] and Monte Carlo methods[45]. GPU computing technology is currently gaining increasing attention from the research community.

This study employs the CRANE code to solve the BEAVRS benchmark. As depicted in Fig. 1, for CRANE, GPU computing is not just for acceleration but the main force of the computation, since nearly all computation-related modules—not only the ones for neutronics calculation but also the ones for thermal-hydraulics and fuel rod calculations—run on GPUs. To resolve the above-mentioned issue of platform availability, the target computing platform of CRANE is chosen as a low-cost mini computer server mounted with consumer-grade GPUs. All calculations for this work are completed on a small server equipped with 10 RTX 3090 graphics cards, and the peak power consumption of the server is 4 kW.

Statistics indicate the average time needed to complete one state multi-physics coupled high-fidelity core analysis is 63.5 s and 101.2 s respectively for Cycle 1 and Cycle 2 of the BEAVRS benchmark. It thus for the first time demonstrates that high-fidelity direct whole core calculation for a large commercial PWR can be completed in about one minute, without resorting to dedicated supercomputers, but merely using low-cost consumer-grade graphics cards. Table 8 summarizes the key computation conditions and corresponding speed performances of multiple direct whole-core calculation codes when applied to the BEAVRS benchmark. It reveals that the superior speed of CRANE is attained despite employing a greater number of energy groups and axial layers compared to other codes.

To better understand the advantages of GPUs in deterministic MOC methods, Fig. 17 [Figure 17: see original paper] gives the computational time breakdown of the CRANE BEAVRS Cycle 1 depletion calculations. It can be seen that the neutronics calculations account for the vast majority of the entire multi-physics coupling computation time. This is straightforward since only the simple closed channel TH module is used for this study. For the neutronics calculations, the 2D planar MOC calculations take the largest portion of computation time. It is no surprise that the 2D planar MOC calculation is the most time-consuming, as it inherently involves the heaviest computational load in the 2D/1D coupled calculation method. Interestingly, the results in the figure show that the 2D MOC calculation accounts for just over 30% of the total computation time. This indicates that the massively parallel architecture of GPUs is exceptionally well-suited for solving the inherently parallel problem of MOC's massive characteristic line tracking, allowing the GPU's computing power to be fully

utilized.

Another advantage of developing a code based on the CPU-GPU heterogeneous platform is the ability to continuously leverage the rapidly advancing GPU technology to enhance the computational speed of the code. The RTX 3090 graphics card used in this study was released in 2020, and since then, NVIDIA has successively released the RTX 4090 and 5090 graphics cards with superior performance. The authors performed a test run of the CRANE code on a platform equipped with RTX 4090 GPUs, and the new computation time is just about 50-60% of that for the current 3090 platform. This implies that the high-fidelity calculation of a single core state of a large commercial reactor, which currently takes 1-2 minutes to complete, can basically be finished within 1 minute on the new 4090 platform. Considering the rapid increase in the performance of GPU graphics cards, the authors believe that these GPU-based high-fidelity direct whole-core calculation codes such as CRANE are likely to become widely adopted as routine tools in the nuclear power industry.

It is noteworthy that although GPUs offer significant advantages for large-scale parallel computing in reactor simulations, they also have inherent limitations that software developers must carefully consider. The most notable disadvantage of GPU computing is its memory constraints, which include both limited VRAM capacity and bandwidth bottlenecks. In terms of VRAM capacity, even the most high-end GPUs today typically have only a few tens of gigabytes of memory (the consumer-grade RTX 3090 used in this study has only 24 GB), whereas CPUs can be paired with terabytes of memory, and supercomputing nodes may offer even more. Regarding bandwidth, although GPU VRAM provides high bandwidth (e.g., the H100 achieves 3 TB/s), data must be transferred via the PCIe bus to CPU memory (with a bandwidth of only about 64 GB/s), which can easily become a bottleneck.

To enable high-fidelity direct 3D full-core calculations on a small server equipped with 10 RTX 3090 GPUs, the CRANE software was developed with extensive optimizations for data storage and communication. Key measures include optimizing MOC calculations by carefully selecting characteristic line densities and discrete angular directions, as well as adopting a modular ray storage approach to significantly reduce the total storage required for ray tracing. Additionally, the burnup calculation is simplified by merging multiple flat-source approximation zones within a fuel pellet into a single burnup region, reducing memory usage for burnup-related data. The software also leverages NVIDIA CUDA Unified Memory technology, which automatically spills over to CPU memory when GPU VRAM is exhausted. Through these optimizations, CRANE can perform high-fidelity criticality and burnup calculations for large commercial PWRs (such as the BEAVRS benchmark's radial 1/4-core model) within the hardware constraints of a single server. Key simulation parameters related to memory usage are as follows: employing the 69-group structure of the cross-section library to perform the neutron transport calculation; tracing the depletion of 100 burnable nuclides; dividing the core into 40 axial layers; and performing the

2D planar MOC calculations with a characteristic ray width of 0.05 cm and 36 discrete angular directions per octant.

VI. Conclusion

In this study, the BEAVRS benchmark is successfully analyzed with the CRANE code, which is a fully GPU-based deterministic direct whole core calculation code targeted for PWR practical applications. Detailed reactor core models are set up, which faithfully represent the practical core of the initial two cycles. Core follow calculations are performed by adopting the exact power history of the reactor. Validation against all the measurement data shows that the predicted criticality, control rod bank worths, in-core detector signal distribution, and the boron let-down curves of two cycles agree well with the measurements. The deviations of the predicted results are all within the acceptance criteria for engineering applications. Moreover, it is demonstrated that by fully exploiting the GPU computing power, direct whole core calculation without resorting to dedicated supercomputers mounted with thousands of cores is possible. Direct whole core calculation with detailed core model can be realized on a low-cost mini computer server mounted with 10 consumer-grade RTX 3090 graphics cards, and the average time needed to complete multi-physics coupled high-fidelity analysis for a single core state point is just about one minute. Thus, it demonstrates that for GPU-based deterministic direct whole-core calculation codes like CRANE, even when utilizing only the computing power of today's consumer-grade GPUs, their accuracy and speed have reached a level sufficient for practical PWR engineering applications. With further advancements in GPU computing technology, codes such as CRANE are anticipated to potentially serve as next-generation routine design tools.

References

- [1] Joo, H.G., Cho, J.Y., Kim, K.S. et al, Methods and Performance of a Three-Dimensional Whole-Core Transport Code DeCART. International Conference on Physics of Reactors (PHYSOR 2004); 2004 Apr 25–29; Chicago, IL.
- [2] Jung, Y.S., Shim, C.B., Lim, C.H. et al, Practical numerical reactor employing direct whole core neutron transport and subchannel thermal/hydraulic solvers. *Ann. Nucl. Energy* 62, 357–374. (2013) DOI:10.1016/j.anucene.2013.06.031
- [3] Kochunas, B., Collins, B., Martin, W.R. et al, Overview of Development and design of MPACT: Michigan parallel characteristics transport code. Intl. Conf. Mathematics and Computational Methods Applied to Nuclear Science & Technology (M&C 2013); 2013 May 5–9; Sun Valley.
- [4] Chen, J., Liu Z., Zhao, C. et al, A new high-fidelity neutronics code NECP-X. *Ann. Nucl. Energy* 116, 417–428.(2018) DOI:10.1016/j.anucene.2018.02.049

- [5] Sutton, T.M., Tj, D., Trumbull, T. et al, The MC21 Monte Carlo Transport Code. Joint International Topical Meeting on Mathematics and Computations and Supercomputing in Nuclear Applications.(2007)
- [6] Wang, K., Li, Z., She, D. et al, RMC – A Monte Carlo code for reactor core analysis. *Ann. Nucl. Energy* 82, 121–129.(2015) DOI:10.1016/j.anucene.2014.08.048
- [7] Romano, P.K., Horelik, N.E., Herman, B.R. et al, OpenMC: A state-of-the-art Monte Carlo code for research and development. *Ann. Nucl. Energy* 82, 90–97.(2014) DOI: 10.1016/j.anucene.2014.07.048
- [8] Deng, L., Ye, T., Li, G. et al, 3-D Monte Carlo Neutron–Photon Transport Code JMCT and Its Algorithms. *PHYSOR 2014*, Kyoto, Japan, September 28–October 3, 2014
- [9] Wu, Y., Song, J., Zheng, H. et al, CAD-based Monte Carlo program for integrated simulation of nuclear system SuperMC. *Ann. Nucl. Energy* 82, 161–168.(2015) DOI:10.1016/j.anucene.2014.08.058
- [10] Horelik, N., Herman, B., Forget, B. et al, Benchmark for evaluation and validation of reactor simulations (BEAVRS). *Intl. Conf. Mathematics and Computational Methods Applied to Nuclear Science & Technology (M&C 2013)*; 2013 May 5–9; Sun Valley.
- [11] Zhang, B., Liu, Z., Yuan, X. et al, Validation of DDC-3D code system for neutronics and thermal-hydraulics coupling analysis using BEAVRS benchmark. *Nucl. Eng. Des.* 429, 113583.(2024) DOI:10.1016/j.nucengdes.2024.113583
- [12] Chen, D., Wu, H., Li, Y. et al, Analysis of BEAVRS Benchmark with CASMO-4E/SIMULATE-3 Code System. *Yuanzineng Kexue Jishu/Atomic Energy Sci. Technol.* 51, 457–461.(2017) (in Chinese) DOI:10.7538/yzk.2017.51.03.0457
- [13] Wang, Z., Wu, B., Hao, L. et al, Validation of SuperMC with BEAVRS benchmark at hot zero power condition. *Ann. Nucl. Energy* 709–714.(2017) DOI:10.1016/j.anucene.2017.09.045
- [14] Yang, W., Liu, B., Yao, S. et al, Calculation of Detector Response Rate Based on BEAVRS Benchmark Using PWR-Core Analysis Code CORAL. *Journal of Nuclear Fuel Cycle and Waste Technology.* DOI:10.1115/ICONE29-93495
- [15] Li, G., Deng, L., Zhang, B. et al, JMCT Monte Carlo analysis of BEAVRS benchmark: hot zero power results. *Acta Phys. Sin. -Chin. Ed.-* 65.(2016) (in Chinese) DOI:10.7498/aps.65.052801
- [16] Leppänen, J., Mattila, R., Pusa, M., Validation of the Serpent-ARES code sequence using the MIT BEAVRS benchmark – Initial core at HZP conditions. *Ann. Nucl. Energy* 69, 212–225.(2014) DOI:10.1016/j.anucene.2014.02.014
- [17] Darnowski, P., Pawluczyk, M., Analysis of the BEAVRS PWR benchmark using SCALE and PARCS. *Nucl. Orig. Ed.* 64, 87–96.(2019)

DOI:10.2478/nuka-2019-0011

[18] Deng, L., Li, G., Zhang, B. et al, JMCT Monte Carlo Simulation Analysis of BEAVRS and SG-III Shielding. EPJ Web Conf 153, 06017.(2017) DOI:10.1051/epjconf/201715306017

[19] Shen, Z., Sun, Q., He, D. et al, Comparison and verification of NECP-X and OpenMC using high-fidelity BEAVRS benchmark models. Nuclear Techniques, 45, 73–81.(2022) (in Chinese) DOI:10.11889/j.0253-3219.2022.hjs.45.010602

[20] Li, Z., Wu, H., Cao, L. et al, On-line monitoring analysis of BEAVRS benchmark using NECP-ONION. PHYSOR 2016, Sun Valley, ID, May 1–5, 2016

[21] Tang, X., Liang, J., Wang, K. et al, Verification on criticality calculation of RMC with BEAVRS full-core benchmark. Nuclear Power Engineering, 35, 235-238.(2015) (in Chinese) DOI:10.13832/j.jnpe.2014.S2.0235

[22] Liu, S., Liang, J., Wu, Q. et al, BEAVRS full core burnup calculation condition power RMC code. Ann. Nucl. Energy 101, 434–446.(2017) DOI:10.1016/j.anucene.2016.11.033

[23] Ryu, M., Jung, Y.S., Cho, H.H. et al, Solution of the BEAVRS benchmark using the nTRACER direct whole core calculation code. J. Nucl. Sci. Technol. 52, 961–969.(2015) DOI:10.1080/00223131.2015.1038664

[24] Yu, J., Lee, H., Kim, H. et al, Simulations of BEAVRS depletion benchmark with MCS/CTF coupling system. Nucl. Eng. Technol. 52, 661–673.(2020) DOI:10.1016/j.net.2019.09.007

[25] Wang, K., Liu, S., Li, Z. et al, Analysis of BEAVRS two-cycle benchmark using RMC based on full core detailed model. Prog. Nucl. Energy 98, 301–312.(2017) <https://doi.org/10.1016/j.pnucene.2017.04.009> DOI:10.1016/j.pnucene.2017.04.009

[26] Park, H.J., Kim, S.J., Kwon, H. et al, BEAVRS benchmark analyses by DeCART stand-alone calculations and comparison with DeCART/MATRA multi-physics coupling calculations. Nucl. Eng. Technol. 52, 1896–1906.(2020) DOI:10.1016/j.net.2020.02.015

[27] Collins, B., Godfrey, A., Stimpson, S. et al, Simulation of the BEAVRS benchmark using VERA. Ann. Nucl. Energy 145, 107602.(2020) DOI:10.1016/j.anucene.2020.107602

[28] Chen, G., Jiang, X., Wang, T. et al, Development of a GPU-based numerical reactor physics code CRANE. 19th Reactor Numerical Computation and Particle Transport Conference, Shanghai, 2023. (in Chinese)

[29] Zhang, S., Chen, G., Development of a new lattice physics code robin for PWR application. Intl. Conf. Mathematics and Computational Methods Applied to Nuclear Science & Technology (M&C 2013); 2013 May 5–9; Sun Valley.

- [30] Salko, R., Wysocki, A., Blyth, T. et al, CTF: A modernized, production-level, thermal hydraulic solver for industry-relevant challenge problems in pressurized water reactors. Nucl. Eng. Des. 397.(2022) DOI:10.1016/j.nucengdes.2022.111927
- [31] Cong, T., Chen, H., Chen, C. et al, Development and evaluation of fuel performance analysis code FuSPAC. Int. J. Adv. Nucl. React. Des. Technol. 4, 129–138.(2022) DOI:10.1016/j.jandt.2022.08.002
- [32] Yamamoto, A., Tabuchi, M., Sugimura, N. et al, Derivation of Optimum Polar Angle Quadrature Set for the Method of Characteristics Based on Approximation Error for the Bickley Function. J. Nucl. Sci. Technol. 44, 129–136.(2007) DOI:10.1080/18811248.2007.9711266
- [33] Rombough, C.T., Adam, P.D., Attard A.C. et al, American National Standard Reload Startup Physics Tests for Pressurized Water Reactors. American Nuclear Society - ANS; La Grange Park (United States); ANSI/ANS-19.6.1-2011.
- [34] Kumar, A., Rathbun, M. S., Liang, J., Core Multi-Physics Integration Benchmark with Measured Data NEUP 14-6742: Final Report. <https://www.osti.gov/servlets/purl/1432668>
- [35] Han, Y., Jiang, X., Wang, D. et al, CMFD and GPU acceleration on method of characteristics for hexagonal cores. Nuclear Engineering and Design 280, 210-222(2014) DOI:10.1016/j.nucengdes.2014.09.038
- [36] Boyd, W., Samuel, S., Li, L. et al, The OpenMOC method of characteristics neutral particle transport code. Ann. Nucl. Energy 68, 43-52(2014) DOI:10.1016/j.anucene.2013.12.012
- [37] Song, P., Zhang, Z., Zhang, Q. et al, Implementation of the CPU/GPU hybrid parallel method of characteristics neutron transport calculation using the heterogeneous cluster with dynamic workload assignment. Ann. Nucl. Energy 135, 106957(2020) DOI:10.1016/j.anucene.2019.106957
- [38] Song, L., Zhang, Q., Zhang, Z. et al, Investigation of the efficiency optimization for the improved subgroup resonance self-shielding treatment on the GPU platform. Ann. Nucl. Energy 159, 108318(2021) DOI:10.1016/j.anucene.2021.108318
- [39] Zou, H., Zhang, Q., Liang, L. et al, Development and verification of a GPU-based MOC transient solver. Ann. Nucl. Energy 174, 109171(2022) DOI:10.1016/j.anucene.2022.109171
- [40] Zou, H., Zhang, Q., Song, P. et al, Development of High-Fidelity Neutron Transport Solver of Alpha. PBNC 2022, November 1-4, Beijing&Chengdu, China
- [41] Choi, N., Kang, J., Lee, H.G. et al, Practical acceleration of direct whole-core calculation employing graphics processing units. Prog. Nucl. Energy 133,

103631.(2021) DOI:10.1016/j.pnucene.2021.103631

[42] Lee, H.G., Kim, K.M., Joo, H.G., Development of scalable GPU-based direct whole-core depletion calculation methods. Prog. Nucl. Energy 165, 104928.(2023) DOI:10.1016/j.pnucene.2023.104928

[43] Zhang, A., Dai, M., Cheng, M. et al, Development of a GPU-based three-dimensional neutron transport code. Ann. Nucl. Energy 109156(2022) DOI:10.1016/j.anucene.2022.109156

[44] Li, S., Wang, Z., Bu, L. et al, ANT-MOC: Scalable Neutral Particle Transport Using 3D Method of Characteristics on Multi-GPU Systems. The International Conference for High Performance Computing 23, November 12, 2023.

[45] Kim, K.M., Im, J., Choi, N. et al, High-performance and high-fidelity GPU-based monte carlo solutions to the BEAVRS benchmark. Nucl. Sci. Eng. 197, 1823–1844.(2023) DOI:10.1080/00295639.2022.2148812

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.